CLAIMS

We claim:

1. A method for storing data, comprising:

distributing a first plurality of groups of logical addresses among one or more storage devices;

receiving a second plurality of data-sets containing the data to be stored;

assigning each data-set among the plurality of data-sets a number chosen from a first plurality of different numbers;

partitioning each data-set into multiple partitions, so that each partition among the multiple partitions receives a sequential partition number;

assigning each partition within each data-set to be stored at a specific group of logical addresses in accordance with the sequential partition number of the partition and the number assigned to the data-set; and

storing each partition at the assigned specific group of logical addresses.

2. The method according to claim 1, wherein the multiple partitions comprise equal size partitions.

3. The method according to claim 1, wherein the data-sets comprise data from at least one of a file, file meta-data, a storage object, a data packet, a video tape, a music track, an image, a database record, contents of a logical unit, and an email.

4. The method according to claim 1, wherein the first plurality of groups comprises s groups each having a different integral group number between 1 and s, wherein the number comprises an integer r randomly chosen from and including integers between 0 and s-1, wherein the sequential partition number comprises a positive integer

p, and wherein the group number of the assigned specific group is (r+p)modulo(s) if (r+p)modulo(s)≠0, and s if (r+p)modulo(s)=0.

5. The method according to claim 1, wherein the one or more storage devices are operative in at least one of a storage area network, a network attached storage system, and an object storage architecture.

6. The method according to claim 1, wherein the number is chosen by a randomizing function.

7. The method according to claim 1, wherein the number is chosen by a consistent hashing function.

8. A method for data distribution, comprising:

receiving at least part of a data-set containing data;

delineating the data into multiple partitions;

distributing logical addresses among an initial set of storage devices so as to provide a balanced access to the devices;

transferring the partitions to the storage devices in accordance with the logical addresses;

adding an additional storage device to the initial set, thus forming an extended set of the storage devices comprising the initial set and the additional storage device; and

redistributing the logical addresses among the storage devices in the extended set so as to cause a portion of the logical addresses and the partitions stored thereat to be transferred from the storage devices in the initial set to the additional storage device, while maintaining the balanced access and without requiring a substantial transfer of the logical addresses among the storage devices in the initial set.

9. The method according to claim 8, wherein the data-

set comprises data from at least one of a file, file meta-data, a storage object, a data packet, a video tape, a music track, an image, a database record, contents of a logical unit, and an email.

10. The method according to claim 8, wherein the initial set of storage devices and the additional storage device are operative in at least one of a storage area network, a network attached storage system, and an object storage architecture.

11. The method according to claim 8, wherein distributing the logical addresses comprises:

generating a first plurality of sets of logical addresses,

and wherein delineating the data comprises:

assigning the at least part of the data-set a number chosen from a first plurality of different numbers; and

assigning each partition among the multiple partitions a sequential partition number,

and wherein transferring the partitions comprises:

storing each partition at one of the sets of logical addresses in accordance with the sequential partition number of the partition and the number.

12. A method for data distribution, comprising:

receiving at least part of a data-set containing data;

delineating the data into multiple partitions;

distributing logical addresses among an initial set of storage devices so as to provide a balanced access to the devices;

transferring the partitions to the storage devices in accordance with the logical addresses;

removing a surplus storage device from the initial set, thus forming a depleted set of the storage devices

comprising the initial set less the surplus storage device; and

redistributing the logical addresses among the storage devices in the depleted set so as to cause the logical addresses of the surplus device and the partitions stored thereat to be transferred to the depleted set, while maintaining the balanced access and without requiring a substantial transfer of the logical addresses among the storage devices in the depleted set.

13. The method according to claim 12, wherein the data-set comprises data from at least one of a file, file meta-data, a storage object, a data packet, a video tape, a music track, an image, a database record, contents of a logical unit, and an email.

14. The method according to claim 12, wherein the initial set of storage devices is operative in at least one of a storage area network, a network attached storage system, and an object storage architecture.

15. The method according to claim 12, wherein distributing the logical addresses comprises:

generating a first plurality of sets of logical addresses,

and wherein delineating the data comprises:

assigning the at least part of the data-set a number chosen from a first plurality of different numbers; and

assigning each partition among the multiple partitions a sequential partition number,

and wherein transferring the partitions comprises:

storing each partition at one of the sets of logical addresses in accordance with the sequential partition number of the partition and the number.

16. A data storage system, comprising:

one or more mass-storage devices, coupled to store

partitions of data at respective first ranges of logical addresses (LAs);

a plurality of interim devices, configured to operate independently of one another, each interim device being assigned a respective second range of the LAs and coupled to receive the partitions of data from and provide the partitions of data to the one or more mass-storage devices having LAs within the respective second range; and

one or more interfaces, which are adapted to receive input/output (IO) requests from host processors, to identify specified partitions of data in response to the IO requests, to convert the IO requests to converted-IO-requests directed to specified LAs in response to the specified partitions of data, and to direct all the converted-IO-requests to the interim device to which the specified LAs are assigned.

17. The storage system according to claim 16, wherein at least one of the mass-storage devices has a slow access time, and wherein at least one of the interim devices has a fast access time.

18. The storage system according to claim 16, wherein the one or more mass-storage devices are coupled to provide a balanced access to the first ranges of LAs.

19. The storage system according to claim 16, wherein the storage system is operative in at least one of a storage area network, a network attached storage system, and an object storage architecture.

20. A data storage system, comprising:

one or more storage devices wherein are distributed a first plurality of groups of logical addresses; and

a processing unit which is adapted to:

receive a second plurality of data-sets containing

the data to be stored,

assign each data-set among the plurality of data-sets a number chosen from a first plurality of different numbers,

partition each data-set into multiple partitions, so that each partition among the multiple partitions receives a sequential partition number,

assign each partition within each data-set to be stored at a specific group of logical addresses in the one or more storage devices in accordance with the sequential partition number of the partition and the number assigned to the data-set, and

store each partition in the one or more storage devices at the assigned specific group of logical addresses.

21. The storage system according to claim 20, wherein the multiple partitions comprise equal size partitions.

22. The storage system according to claim 20, wherein the data-sets comprise data from at least one of a file, file meta-data, a storage object, a data packet, a video tape, a music track, an image, a database record, contents of a logical unit, and an email.

23. The storage system according to claim 20, wherein the first plurality of groups comprises s groups each having a different integral group number between 1 and s, wherein the number comprises an integer r randomly chosen from and including integers between 0 and s-1, wherein the sequential partition number comprises a positive integer p, and wherein the group number of the assigned specific group is (r+p)modulo(s) if (r+p)modulo(s)$\neq$0, and s if (r+p)modulo(s)=0.

24. The storage system according to claim 20, wherein the one or more storage devices and the processing unit

are operative in at least one of a storage area network, a network attached storage system, and an object storage architecture.

25. The storage system according to claim 20, wherein the number is chosen by a randomizing function.

26. The storage system according to claim 20, wherein the number is chosen by a consistent hashing function.

27. Data distribution apparatus, comprising:

an initial set of storage devices among which are distributed logical addresses so as to provide a balanced access to the devices;

an additional storage device to the initial set, thus forming an extended set of the storage devices comprising the initial set and the additional storage device; and

a processor which is adapted to receive at least part of a data-set containing data, to delineate the data into multiple partitions, to transfer the partitions to the initial set of storage devices in accordance with the logical addresses, to redistribute the logical addresses among the storage devices in the extended set so as to cause a portion of the logical addresses and the partitions stored thereat to be transferred from the storage devices in the initial set to the additional storage device, while maintaining the balanced access and without requiring a substantial transfer of the logical addresses among the storage devices in the initial set.

28. The apparatus according to claim 27, wherein the data-set comprises data from at least one of a file, file meta-data, a storage object, a data packet, a video tape, a music track, an image, a database record, contents of a logical unit, and an email.

29. The apparatus according to claim 27, wherein the

initial set of storage devices and the additional storage device are operative in at least one of a storage area network, a network attached storage system, and an object storage architecture.

30. The apparatus according to claim 27, wherein the logical addresses comprise a plurality of sets of logical addresses, and wherein the processor is adapted to:

assign the at least part of the data-set a number chosen from a plurality of different numbers,

assign each partition among the multiple partitions a sequential partition number, and

store each partition at one of the sets of logical addresses in accordance with the sequential partition number of the partition and the number.

31. Data distribution apparatus, comprising:

an initial set of storage devices among which are distributed logical addresses so as to provide a balanced access to the devices;

a depleted set of storage devices, formed by subtracting a surplus storage device from the initial set; and

a processor which is adapted to receive at least part of a data-set containing data, to delineate the data into multiple partitions, to transfer the partitions to the initial set of storage devices in accordance with the logical addresses, to redistribute the logical addresses and the partitions stored thereat of the surplus storage device among the storage devices in the depleted set while maintaining the balanced access and without requiring a substantial transfer of the logical addresses among the storage devices in the depleted set.

32. The apparatus according to claim 31, wherein the data-set comprises data from at least one of a file, file

meta-data, a storage object, a data packet, a video tape, a music track, an image, a database record, contents of a logical unit, and an email.

33. The apparatus according to claim 31, wherein the initial set of storage devices is operative in at least one of a storage area network, a network attached storage system, and an object storage architecture.

34. The apparatus according to claim 31, wherein the logical addresses comprise a plurality of sets of logical addresses, and wherein the processor is adapted to:

assign the at least part of the data-set a number chosen from a plurality of different numbers,

assign each partition among the multiple partitions a sequential partition number, and

store each partition at one of the sets of logical addresses in accordance with the sequential partition number of the partition and the number.

35. A method for storing data, comprising:

coupling one or more mass-storage devices to store partitions of data at respective first ranges of logical addresses (LAs);

configuring a plurality of interim devices to operate independently of one another;

assigning each interim device a respective second range of the LAs;

coupling each interim device to receive the partitions of data from and provide the partitions of data to the one or more mass-storage devices having LAs within the respective second range;

receiving input/output (IO) requests from host processors;

identifying specified partitions of data in response to the IO requests;

converting the IO requests to converted-IO-requests directed to specified LAs in response to the specified partitions of data; and

directing all the converted-IO-requests to the interim device to which the specified LAs are assigned.

36. The method according to claim 35, wherein at least one of the mass-storage devices has a slow access time, and wherein at least one of the interim devices has a fast access time.

37. The method according to claim 35, wherein the one or more mass-storage devices are coupled to provide a balanced access to the first ranges of LAs.

38. The method according to claim 35, wherein the one or more storage devices and the plurality of interim devices are operative in at least one of a storage area network, a network attached storage system, and an object storage architecture.

39. A method for data distribution, comprising:

receiving at least part of a data-set containing data;

delineating the data into multiple equal size partitions;

transferring the partitions to an initial set of storage devices so as to provide a balanced access to the devices;

adding an additional storage device to the initial set, thus forming an extended set of the storage devices comprising the initial set and the additional storage device; and

redistributing the partitions among the storage devices in the extended set so as to cause a portion of the partitions to be transferred from the storage devices in the initial set to the additional storage device,

while maintaining the balanced access and without requiring a substantial transfer of the partitions among the storage devices in the initial set.

40. A method for data distribution, comprising:

receiving at least part of a data-set containing data;

delineating the data into multiple equal size partitions;

transferring the partitions to an initial set of storage devices so as to provide a balanced access to the devices;

removing a surplus storage device from the initial set, thus forming a depleted set of the storage devices comprising the initial set less the surplus storage device; and

redistributing the partitions stored in the surplus device to the depleted set, while maintaining the balanced access and without requiring a substantial transfer of the partitions among the storage devices in the depleted set.

41. Data distribution apparatus, comprising:

an initial set of storage devices;

an additional storage device to the initial set, thus forming an extended set of the storage devices comprising the initial set and the additional storage device; and

a processor which is adapted to receive at least part of a data-set containing data, to delineate the data into multiple equal size partitions, to transfer the partitions to the initial set of storage devices so as to provide a balanced access to the initial set of storage devices, to redistribute the partitions among the storage devices in the extended set so as to cause a portion the partitions stored in the initial set to be transferred to

the additional storage device, while maintaining the balanced access and without requiring a substantial transfer of the partitions among the storage devices in the initial set.

42. Data distribution apparatus, comprising:

an initial set of storage devices;

a depleted set of storage devices, formed by subtracting a surplus storage device from the initial set; and

a processor which is adapted to receive at least part of a data-set containing data, to delineate the data into multiple equal size partitions, to transfer the partitions to the initial set of storage devices so as to provide a balanced access to the initial set of storage devices, to redistribute the partitions of the surplus storage device among the storage devices in the depleted set while maintaining the balanced access and without requiring a substantial transfer of the partitions among the storage devices in the depleted set.